

Chapter 19

Volume Organization and Naming

The *Volume Organization and Naming* Standard defines the standard way of organizing data sets onto physical media and the conventions for forming volume names and identifiers. A volume is one unit of physical media such as a CD-ROM, a CD-WO, an 8mm magnetic tape, or a 9-track magnetic tape. Data sets may reside on one or more volumes and multiple data sets may also be stored on a single volume. Volumes are grouped into Volume Sets.

Each volume has a directory structure which contains subdirectories and files. Both random access (CD-ROM) and sequential access (magnetic tape) media are supported. A PDS volume on sequential access media has a “virtual” directory structure defined in the volume object included on the volume in the file VOLDESC.CAT. The virtual directory structure may be used to recreate the volume directory structure when the files are moved to random access media.

PDS recommends that archive volumes be based on a single version of the PDS Standards. Software tools that work with one version of the standard may not work with all versions.

19.1 Volume Set Types

Data may be organized into one of four types of archive volumes. The distinguishing characteristics between the volume types are the number of data sets on each volume and the number of volumes required to capture all the data. The directory organization of the volumes and the required files varies slightly depending on the volume type. Figures 19.1 through 19.5 depict the various volume directory structure options. The four volume types are described below.

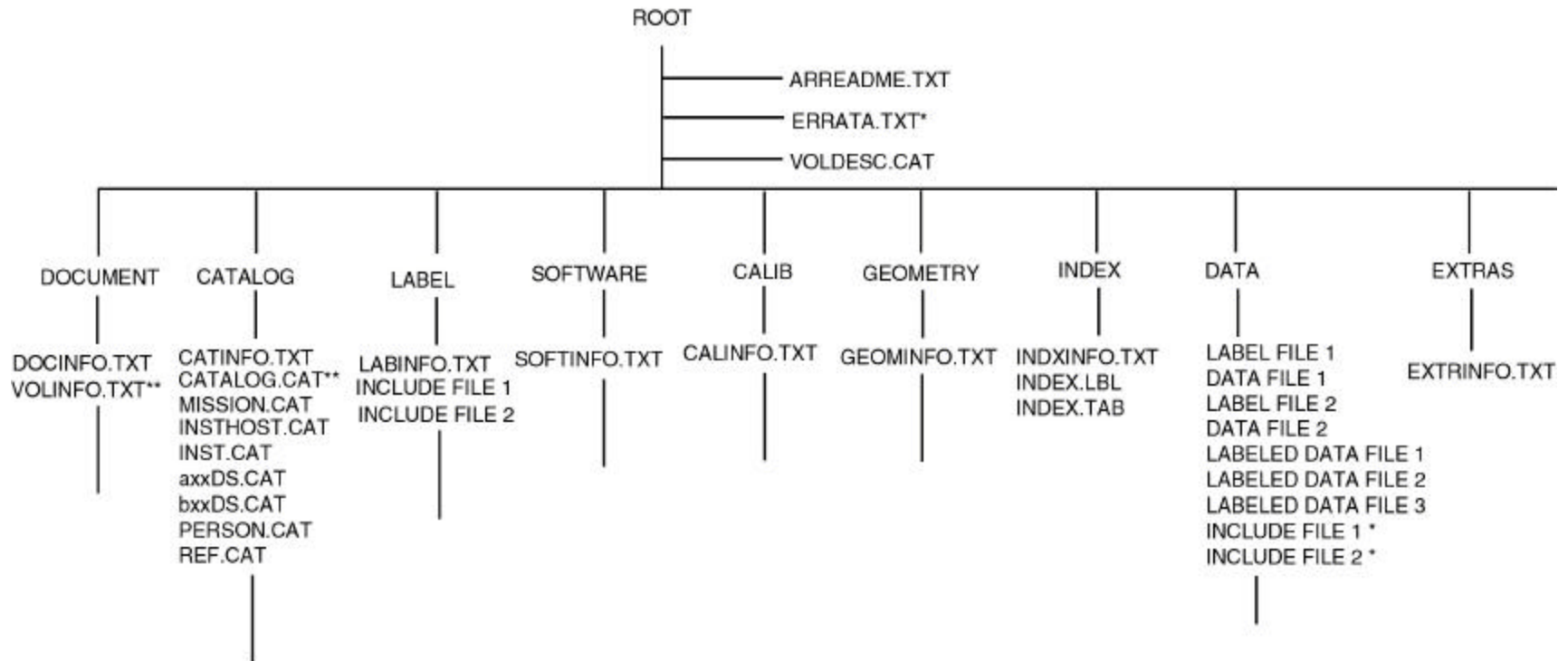
- (1) One data set on one volume - this is the basic volume organization consisting of the required ROOT directory, INDEX, and data subdirectories and the seven optional subdirectories: DOCUMENT, CATALOG, LABEL, GAZETTER (not shown in the figures), SOFTWARE, CALIB, and GEOMETRY. See Figure 19.1.
Note that CALIB and GEOMETRY are only recommended directory names, other appropriate names may be substituted.
- (2) One data set on many volumes - this type includes both an index for the volume and a cumulative index for the volume set (up to the given volume number, not the entire set) in the INDEX subdirectory. See Figure 19.2.
- (3a) Many data sets on one volume (one logical volume) - this type of volume requires additional file naming conventions to distinguish similar files for different data sets. In addition, the DATA subdirectories are organized by data set (or equivalent, e.g. instrument)

at the first level below the ROOT directory. See Figure 19.3.

- (3b) Many data sets on one volume (many logical volumes) - this volume organization is designed to accommodate many small data sets that have distinct documentation, indexing and other ancillary information that are more logically packaged together below the root directory of the volume. See Figure 19.4. Directories common to all logical volumes (e.g. SOFTWARE) may also be supplied, provided there are no pointer references to any files within a common directory.
- (4) Many data sets on many volumes - this type requires additional file naming conventions, cumulative indices, and a first level subdirectory organization by data set. See Figure 19.5.

NOTE: It is permissible to have one or more data volumes with an ancillary volume containing the DOCUMENT, CATALOG, GAZETTER, SOFTWARE, CALIB, and GEOMETRY directories. If this is done, PDS requires that all include files be present on each data disk. PDS prefers that ancillary files be archived on the same volumes as the data wherever possible. This makes data easier to access for the science users. The contents and organization of the directories of all the volume types are described in this chapter.

VOLUME SET ORGANIZATION STANDARD ONE DATA SET, ONE VOLUME



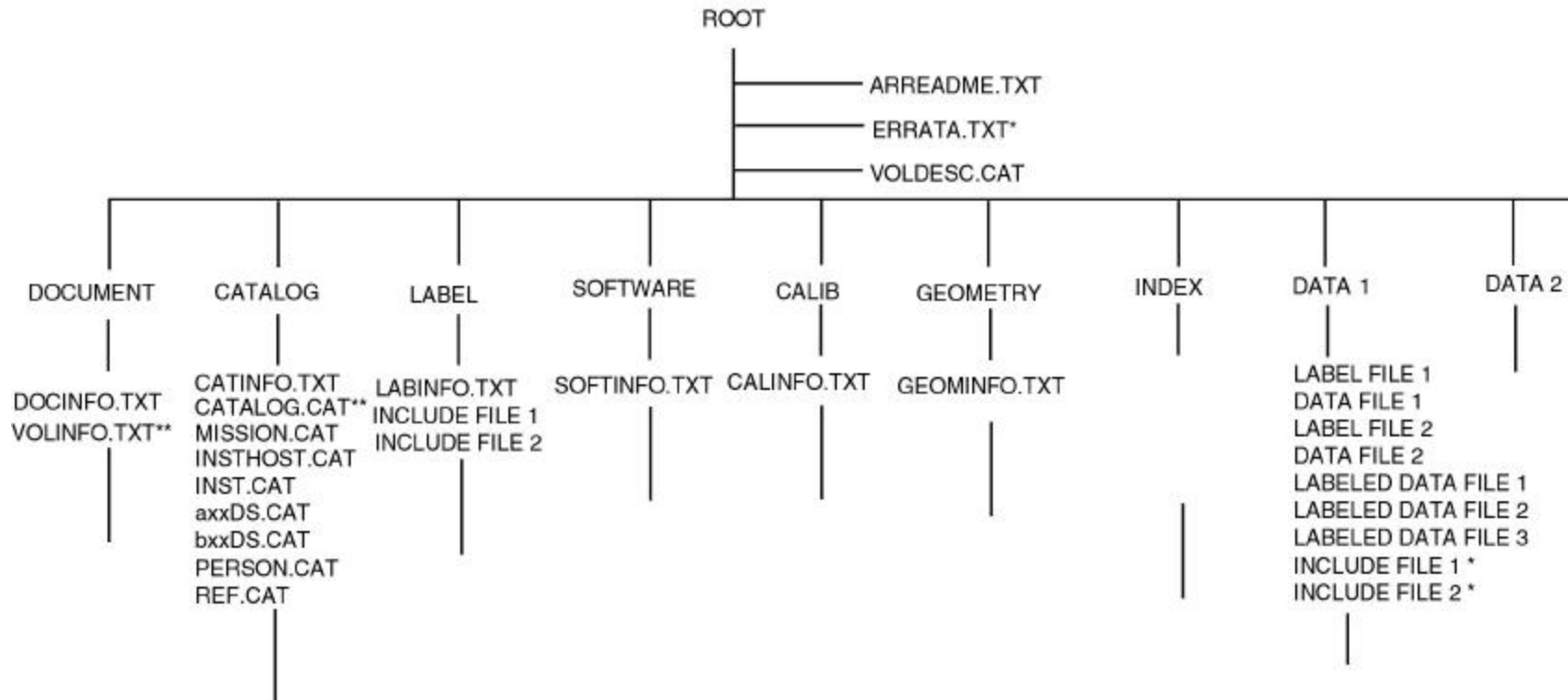
xxxxlINFO.TXT Required for each non-data subdirectory if present

* Optional

** Either catalog objects, or CATALOG.CAT, or VOLINFO.TXT required (listed preferentially)

Figure 19.1 Volume Set Organization Standard - One Data Set, One Volume

VOLUME SET ORGANIZATION STANDARD ONE DATA SET, MANY VOLUMES



xxxxINFO.TXT Required for each non-data subdirectory if present

* Optional

** Either catalog objects, or CATALOG.CAT, or VOLINFO.TXT required (listed preferentially)

Figure 19.2 Volume Set Organization Standard - One Data Set, Many Volumes

VOLUME SET ORGANIZATION STANDARD MANY DATA SETS, ONE VOLUME

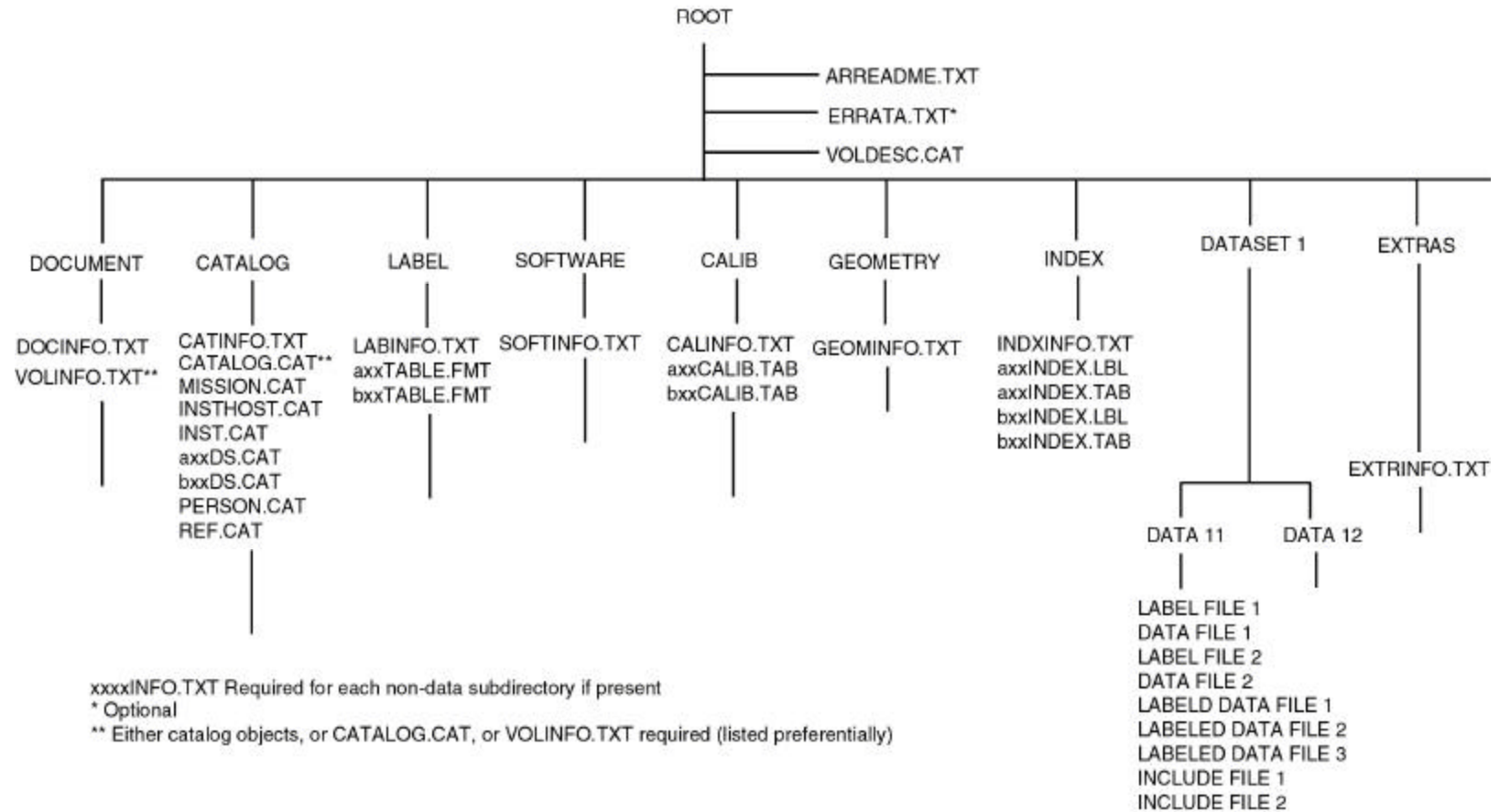
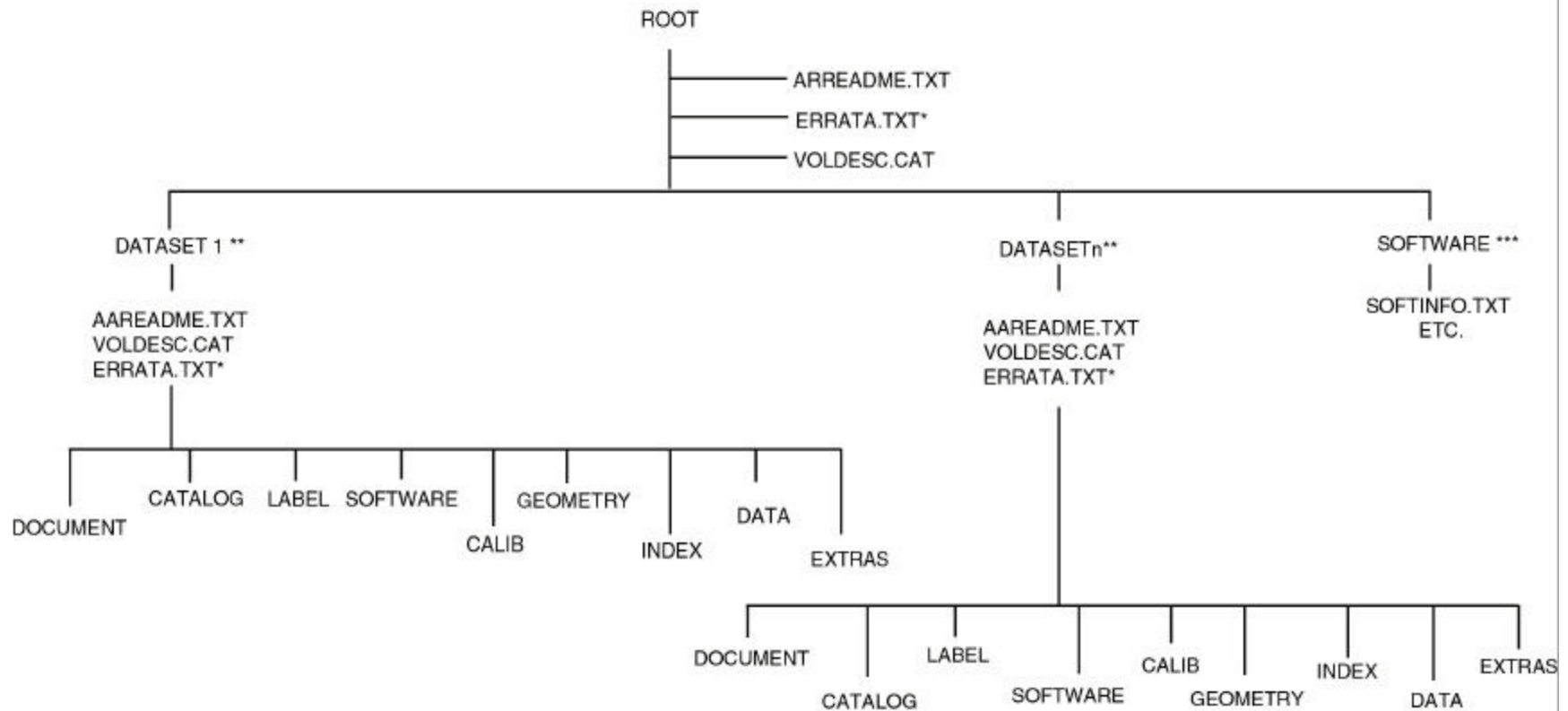


Figure 19.3 Volume Set Organization Standard - Many Data Sets, One Volume

VOLUME SET ORGANIZATION STANDARD
MANY DATA SETS, ONE PHYSICAL VOLUME,
MANY LOGICAL VOLUMES



* Optional

** Logical volume; directory structure identical to Figure 19.1, ONE DATA SET, ONE VOLUME

*** Common to all logical volumes

Figure 19.4 Volume Set Organization Standard - Many Data Sets, One Physical Volume, Many Logical Volumes

VOLUME SET ORGANIZATION STANDARD MANY DATA SETS, MANY VOLUMES

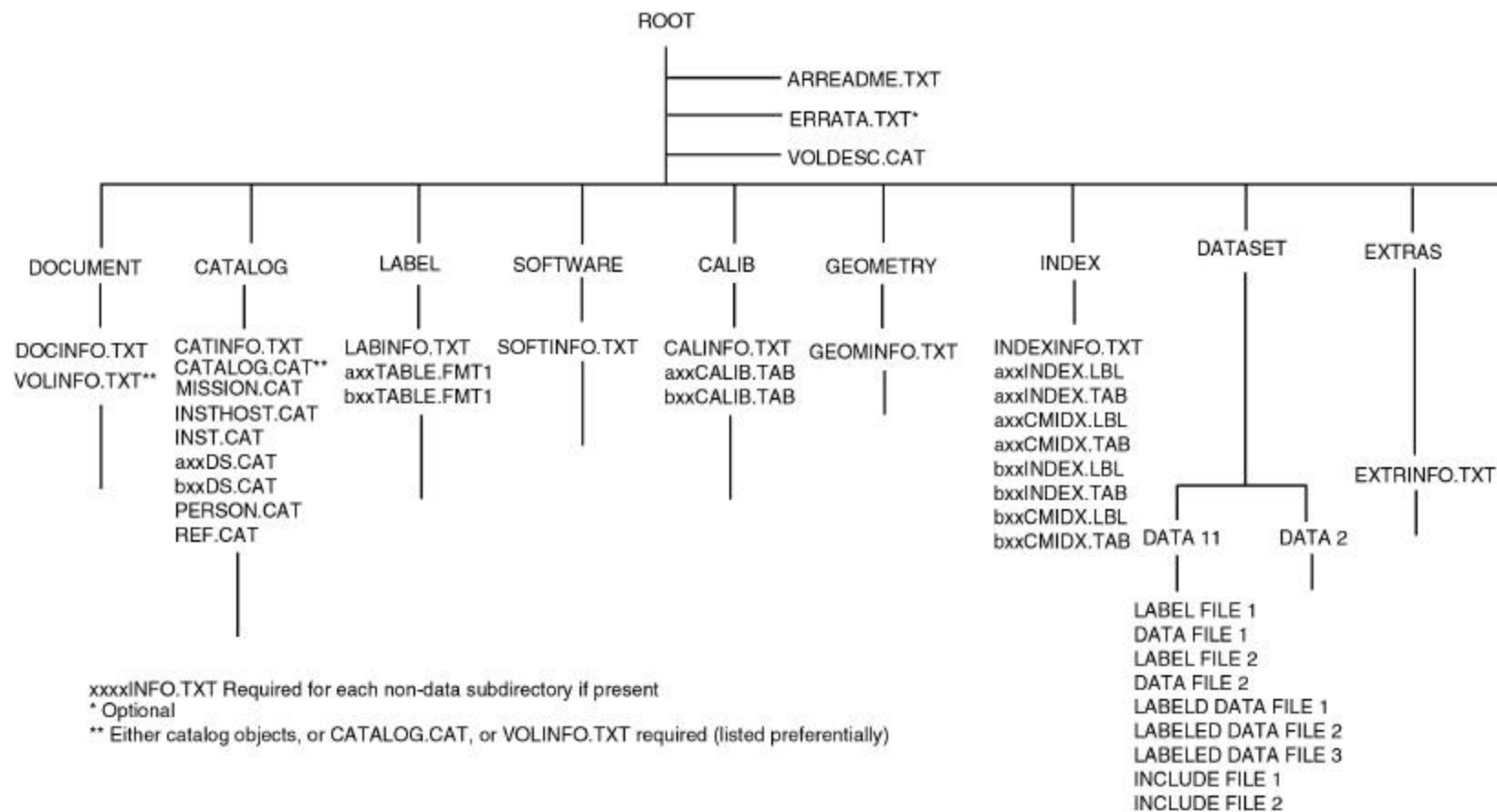


Figure 19.5 Volume Set Organization Standard - Many Data Sets, Many Volumes

19.2 Volume Organization Guidelines

PDS recommends that directory structures be simple, path names short, and directory and file names be constructed in a logical manner. It is recommended that the number of files per subdirectory should ideally be a screenful, allowing users to browse through file names using the directory command. Some externally developed software cannot handle subdirectories with more than 255 files, so it is recommended that this number not be exceeded. PDS also recommends that there be no empty subdirectories (as a convenience to users).

19.3 Description of Directory Contents and Organization

ROOT Directory -- Required

Top level directory of a physical or logical volume. The ROOT directory (of a physical or logical volume) contains the following required and optional files and subdirectories.

AAREADME.TXT -- Required

Contains an overview of the contents of the volume (physical or logical volume) and its organization, general instructions for using the volume and its contents, and provides contact information. Its name has been chosen so that it will be listed first in an alphabetical directory listing. See Appendix D for an outline and example of an AAREADME.TXT file.

ERRATA.TXT -- Optional

Contains textual information describing errors and/or anomalies found in the current volume as well as errors and/or anomalies found in previous volumes of a volume set. If known errors exist on a volume they shall be documented in this file.

VOLDESC.CAT -- Required

Contains the VOLUME Object which gives a high-level description of the contents of the volume.

VOLDESC.SFD -- Optional

Contains the SFDU Reference Object structure which aggregates the separate file contents of the volume into an SFDU. The Reference Object is expressed in PVL. This file should only be considered for use if the data products are packaged as SFDUs. Note: the “SFD” file extension is a reserved file extension in the CCSDS SFDU standard indicating the file contains a valid SFDU. Note that this file is identified here for backward compatibility with previous versions of the PDS standards and is not to be used in current archive products

DOCUMENT Subdirectory -- Optional

Contains all the textual material that describes the mission, spacecraft, instrument, and data set. This can include references to science papers, or the actual papers.

DOCINFO.TXT -- Required

Contains a textual description of the contents of the DOCUMENT subdirectory.

VOLINFO.TXT -- Optional

Contains a textual description of the contents of the volume. It is an optional file, however, either one or both of the VOLINFO.TXT or the data set catalog objects in the CATALOG subdirectory shall be included on the volume (see the CATALOG subdirectory).

CATALOG Subdirectory -- Optional

Contains all the completed catalog objects for the mission(s), instrument host(s), instrument(s), and data set(s) for the archive volume. This is an optional directory (i.e., a complete set of catalog objects do not have to be included on the archive volume); if and only if, a VOLINFO.TXT file is included in the DOCUMENT directory of the volume. If a complete set of catalog objects are provided then the VOLINFO.TXT is not required, and vice versa. The VOLINFO.TXT file is the textual equivalent of the catalog objects (i.e., all of the information required to be present in the catalog objects shall be present in the VOLINFO.TXT file).

The primary difference between catalog objects and the VOLINFO.TXT file is that the catalog objects use an "OBJECT / END_OBJECT" and "KEYWORD = VALUE" format and each catalog object is an independent file. The VOLINFO.TXT is a single file which uses a text format.

Note that for logical volumes, these must be below the logical volume root, if present.

CATINFO.TXT -- Required

Contains a textual description of the contents of the CATALOG subdirectory.

CATALOG.CAT -- Required

Contains the entire set of high-level descriptive information about a data set (this includes separate descriptions for each mission, each instrument host, each instrument, and each data set; as well as, reference and personnel information), expressed in PDS objects which makes the file suitable for loading into a catalog.

PDS Preferred Method for Supplying Catalog Objects

Individual catalog objects may also be packaged into separate files. This is the preferred method for supplying catalog objects as each catalog object must be ingested into the PDS catalog independent of the other objects. If a CATALOG.CAT file is supplied, the CN data engineer must disassemble the single file into multiple files / catalog templates. Each file corresponds to a catalog template and each catalog template corresponds to a single catalog object. The data engineer then must validate each file, and format each description field in accordance with the prescribed headings and sub-headings for each catalog template.

For example, in Figure 19.5, the files axxxxxDS.CAT and bxxxxxDS.CAT represent two separate files each containing data set catalog objects (descriptive information about the data set) for data sets a and b respectively. See the *File Specification and Naming* chapter in this document for the file naming rules. See also Appendix A for the required contents of the catalog object.

Note that the axx- and bxx- prefixes in the sample names are neither required nor recommended. Data producers may use them to distinguish two or more files (by data set, instrument, or other criterion). The data producer should replace the generic prefixes shown here by a suitable mnemonic acronym.

LABEL Subdirectory -- Optional

Contains additional PDS labels and/or include files (meta data or descriptive information) which were not packaged with the data products or in the data subdirectories.

Note that if a logical volume organization is used, the LABEL subdirectory, if present, must reside below the logical volume ROOT, since pointer references to files within a common directory are not allowed.

LABINFO.TXT -- Required

Contains a textual description of the contents of the LABEL subdirectory.

Include Files -- Required

Files pointed to in a PDS label that contain additional meta data or descriptive information. Only files of type LBL, TXT, or FMT shall be included in the LABEL subdirectory. In the figures, the files axxINCLUDE FILE₁, bxxINCLUDE FILE₁ and INCLUDE FILE₁ represent sample files of the above types. The axx and bxx prefixes indicate that the include files for different data sets (a and b) may be combined in the same LABEL subdirectory.

Note that the axx- and bxx- prefixes in the sample names are neither required nor recommended. Data producers may use them to distinguish two or more files (by data set, instrument, or other criterion). The data producer should replace the generic prefixes shown here by a suitable mnemonic acronym.

GAZETTER Subdirectory -- Optional

Contains detailed information about all the named features on a target body associated with the data sets on the volumes. The features are those the International Astronomical Union (IAU) has named and approved.

GAZINFO.TXT -- Required

Contains a textual description of the contents of the GAZETTER subdirectory.

GAZETTER.TXT -- Required

Contains a textual description of the structure and contents of the gazetteer table.

GAZETTER.LBL -- Required

Contains the PDS label identifying and giving a formal description of the structure of the gazetteer table.

GAZETTER.TAB -- Required

Contains the gazetteer table.

SOFTWARE Subdirectory -- Optional

Contains the software libraries, utilities, or application programs to access/process the data objects. It may also include algorithms. Currently only public domain software can be included on PDS archive volumes.

The following SOFTWARE subdirectory structure is the recommended platform-based model. An alternative model for the SOFTWARE subdirectory structure is application-based (e.g. directory names are based on the application such as DISPLAY). See Appendix D SOFTINFO.TXT example for the subdirectory structure used for Clementine. See Appendix E for the subdirectory structure of the NAIF Toolkit for a single platform.

SOFTINFO.TXT -- Required

Contains a textual description of the contents of the SOFTWARE subdirectory.
For an outline and example, see Appendix D.

SRC Subdirectory -- Optional

There can be a global SRC directory under the SOFTWARE directory if there is source code applicable to all platforms. For example, application programming languages such as IDL are relatively platform independent and would be placed in a global SRC directory. Note in example below, there is both a global source directory as well as source directories at the lower levels.

DOC Subdirectory -- Optional

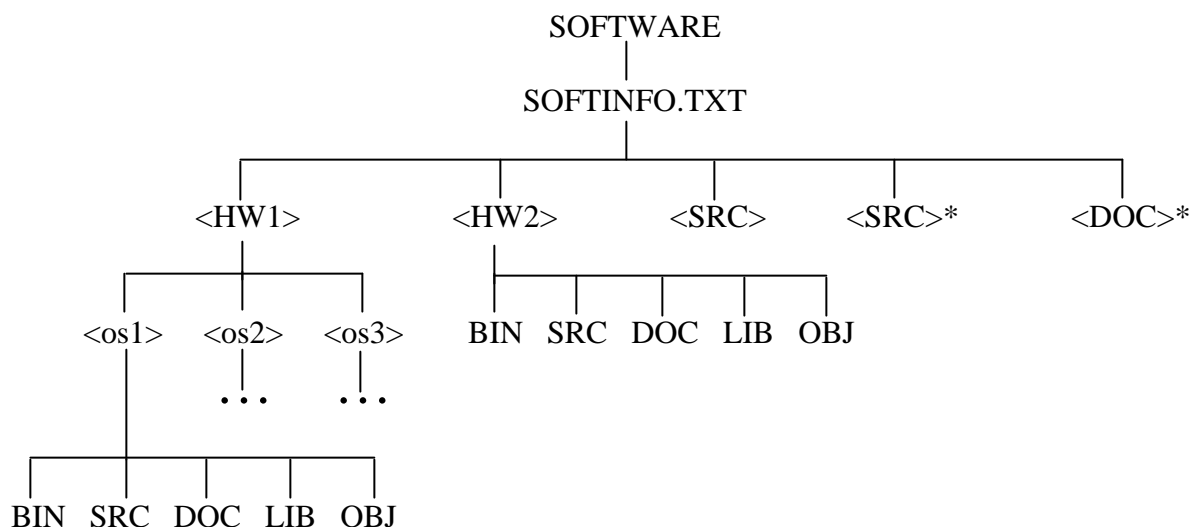
A global DOC directory under the SOFTWARE directory would contain documentation for the source code in the global SRC directory.

LIB Subdirectory -- Optional

A global LIB directory under the SOFTWARE directory would contain libraries applicable to all platforms.

Hardware Platform and Operating System/Environment Subdirectories -- Optional
(not present if only global source code provided)

1. The hardware platform and the operating system/environment must be explicitly stated. If there is more than one operating system/environment (os/env) supported then they must be subdirectories under the hardware directories. If there is only one, then that subdirectory can be promoted to the hardware directory level (via naming conventions). In the example below, since only one os/env is supported on hardware 2, the name of the hardware subdirectory also contains the os/env name.



2. The next level of directories are BIN, SRC, DOC, LIB and OBJ. If any are not applicable, they should be left off (i.e. no empty directories).

*info.txt files under SOFTWARE subdirectories are optional (e.g. PCINFO.TXT, MACINFO.TXT, VAXINFO.TXT, SUNINFO.TXT, etc.).

3. Examples of subdirectory names for the two cases where there are single or multiple operating system/environments are listed below. This list is not meant to be a complete list, it will be updated on an as-needed basis.

Multiple	Single
PC	
DOS	PCDOS
WIN	PCWIN
WINNT	PCWINNT
OS2	PCOS2
MAC	
SYS7	MACSYS7
AUX	MACAUX
SUN	
SUNOS	SUNOS
SOLAR	SUNSOLAR

VAX	
VMS	VAXVMS
ULTRX	VAXULTRX
SGI	
IRX4	SGIIRX4
IRX5	SGIIRX5

CALIBration Subdirectory -- Optional

Contains the calibration files used in the processing of the raw data or needed to use the data products on the volume.

Note that CALIB is only a recommended directory name, another appropriate name may be used.

CALINFO.TXT -- Required

Contains a textual description of the contents of the CALIB subdirectory.

Calibration Files -- Required

In the figures, the files axxCALIB.TAB and bxxCALIB.TAB represent sample files. The axx and bxx prefixes indicate that the calibration files for different data sets (a and b) may be combined in the same CALIB subdirectory.

Note that the axx- and bxx- prefixes in the sample names are neither required nor recommended. Data producers may use them to distinguish two or more files (by data set, instrument, or other criterion). The data producer should replace the generic prefixes shown here by a suitable mnemonic acronym.

GEOMETRY Subdirectory -- Optional

Contains the relevant files (e.g., SEDRs, SPICE kernels) needed to describe the observation geometry.

Note that GEOMETRY is only a recommended directory name, another appropriate name may be used.

GEOMINFO.TXT -- Required

Contains a textual description of the contents of the GEOMETRY subdirectory.

INDEX Subdirectory -- Required (exception noted below)

Contains the indices for the data products in the data set(s) on the volume.

Exception note: If the logical volume organization is used, there will generally be no INDEX subdirectory at the ROOT of the physical volume. Instead there will be individual INDEX subdirectories at the ROOT of each logical volume.

INDXINFO.TXT -- Required

Contains a textual description of the contents of the INDEX subdirectory. This description should include at least:

- 1) A description of the structure and contents of each index table in this subdirectory.
- 2) Usage notes

For an example of the INDXINFO.TXT file, see Appendix D, Section D.2.

INDEX.LBL -- Required (exception noted below)

For all volumes, this file contains the PDS label for the volume index (INDEX.TAB). The INDEX_TABLE specific object should be used to identify and describe the structure (columns) of the index table. See Appendix A.

Although INDEX.LBL is the preferred name for this file, the name axxINDEX.LBL may also be used (with axx replaced by an appropriate mnemonic).

Exception note: PDS recommends the use of detached labels for index tables. If an attached label is used, this file is superfluous (i.e., not needed).

INDEX.TAB -- Required

For all volumes, this file contains the volume index in tabular format. Normally only data files are included in an index table. In some cases, however, ancillary files may be included.

Although INDEX.TAB is the preferred name for this file, the name axxINDEX.TAB may also be used (with axx replaced by an appropriate mnemonic).

Note that the axx- and bxx- prefixes in the sample names are neither required nor recommended. Data producers may use them to distinguish two or more files (by data set, instrument, or other criterion). The data producer should replace the generic prefixes shown here by a suitable mnemonic acronym.

CUMINDEX.LBL -- Recommended for multi-volume sets

For multi-volume sets, this file contains the PDS label for the cumulative volume set index (CUMINDEX.TAB). The INDEX_TABLE specific object should be used to identify and describe the structure (columns) of the cumulative volume set index table. See Appendix A. Although CUMINDEX.LBL is the preferred name for this file, the name axxCUMIDX.LBL may also be used (with axx replaced by an appropriate mnemonic).

PDS recommends the use of detached labels for index tables. If an attached label is used, this file is not needed.

CUMINDEX.TAB --Recommended for multi-volume sets

For multi-volume sets, this file contains the cumulative volume set index in a tabular format. Normally only data files are included in a cumulative index table. In some cases, however, ancillary files may be included.

Although CUMINDEX.TAB is the preferred name for this file, the name axxCUMIDX.TAB may also be used (with axx replaced by an appropriate mnemonic).

EXTRAS Subdirectory -- Optional

The EXTRAS directory is the designated area for housing additional elements provided by data

producers beyond the scope of PDS compliance requirements of a data set. Examples include HTML-based disk navigators, educational and public interest aids, and other useful but nonessential items. The PDS has no restrictions on the contents and organization of this subdirectory other than conformance to ISO-9660/UDF standards.

EXTRINFO.TXT -- Required

Contains a textual description of the contents and organization of the EXTRAS subdirectory. This description should include at least:

- 1) A description of the structure and contents of each file in this subdirectory.
- 2) Usage notes

Data Subdirectories -- Required (exception noted below)

Contain the data product files. These subdirectories are organized and named according to the *Directory Types and Naming* chapter in this document. Subdirectories may be nested up to eight levels deep on a physical volume. Data products may be packaged with their PDS labels attached, where the label and the data object(s) are contained in a LABELED DATA FILE, or with PDS labels detached, where the PDS label is contained in a LABEL FILE and the data object(s) in a DATA FILE.

Data File -- Contains a data object which is a grouping of data resulting from a scientific observation such as an image or table, representing the measured instrument parameters. The associated PDS label is contained in a LABEL FILE.

Label File -- Contains a detached PDS label expressed in the Object Definition Language that identifies, describes, and defines the structure of the data objects. The associated data objects are contained in a DATA FILE. The LABEL FILE shall have the same basename as the associated DATA FILE and the extension of ".LBL".

Labeled Data File -- Contains data object(s) and associated PDS label.

Exception note: Data subdirectories are not present at the ROOT level of a physical volume when logical volumes are used. Instead, they are nested below the ROOT of the logical volume.

19.4 Volume Naming

The Volume name provides the name of a data volume. Volume names shall be at most 60 characters in length and are in upper case. They should describe the contents of the volume in terms that a human user can understand. Most computer systems and software use the volume ID, not the volume set name or volume name, when processing media volumes. The volume set name or volume name are therefore more important to a human user than to a machine.

In most cases the volume name is more specific than the volume set name. For example, the

volume name for the first volume in the VOYAGER IMAGES OF URANUS volume set is:

"VOLUME 1: COMPRESSED IMAGES 24476.54 - 26439.58"

19.4.1 Volume ID

Many types of media and the machines that read media volumes place a limit on the length of the volume ID. Therefore, although the complete volume set ID should be placed on the outside label of the volume, a shorter version is actually used when the volume is recorded. PDS has adopted a limit of 9 characters for these terse volume identifiers. This terse identifier shall consist of the last two components of the volume set ID, with the "X" wildcard values replaced by the sequence number associated with the particular volume (see the *Volume Set ID* Standard below). This ID must always be unique for PDS data volumes. Note that the ID must be in upper case.

EXAMPLES:

VG_0002 (for volume 2 of the Voyager set)

MG_0001 (for the first volume of the Magellan set)

VGRS_0001 (for a potential Voyager Radio Science collection)

If a volume is redone because of errors in the initial production the volume ID should remain the same, and the VOLUME_VERSION_ID should be incremented. This parameter is contained in the VOLDESC.CAT file on the volume, and the version ID should also be placed on the external volume label as "Version n" where n indicates the revision number. This indicates that the original volume should be replaced with the new version. If a volume is redone because the data have been enhanced it should be given a new volume ID, not a new version number.

19.5 Volume Set Naming

The Volume Set Name provides the full, formal name of a group of data volumes containing a data set or a collection of related data sets. Volume set names shall be at most 60 characters in length and must be in upper case. Volume sets are normally considered as a single orderable entity.

For example, the volume series MISSION TO VENUS consists of the following volume sets:

MAGELLAN: THE MOSAIC IMAGE DATA RECORD

MAGELLAN: THE ALTIMETRY AND RADIOMETRY DATA RECORD

MAGELLAN: THE GLOBAL ALTIMETRY AND RADIOMETRY DATA RECORD

PRE-MAGELLAN RADAR AND GRAVITY DATA SET COLLECTION

In certain cases, the volume set name can be the same as the volume name, such as when the volume set consists of only one volume.

Note that in VAX computer usage a volume set has very special attributes, and that all volumes of a volume set must be on line for proper access. There are no plans within PDS to produce volume sets following the VAX definition. Instead the VOLUME_SET_NAME and

VOLUME_SET_ID are used to group related data and to provide additional specificity in a volume name in case volumes produced by different organizations have the same volume IDs.

19.5.1 Volume Set ID

The volume set ID identifies a data volume or a set of volumes. Volume sets are normally considered as a single orderable entity. Volume set IDs shall be at most 60 characters in length, must be in upper case, and are formed of the following fields, separated by underscores:

1. The country of origin (abbreviated).
2. The government branch.
3. The discipline within the branch that is producing the volumes.
4. A campaign, mission or spacecraft identifier (2 characters) followed by an optional 2 character instrument or product identifier.
5. A 4 digit sequence identifier. The first digit or digits may be used to represent the volume set and the trailing "X"s are wildcards that represent the range of volumes in the set. Up to 4 "X"s are allowed.

EXAMPLE

USA_NASA_PDS_GO_10XX could be the Volume set ID for the Galileo EDR volume set, since there are less than 100 volumes (since the XX placeholder accommodates the range 01 - 99 only). Note that the volume IDs for volumes in the set would then be GO_1001, GO_1002, etc.

NOTE: Prior to version 3.2, the 4-digit sequence identifier (item 5 above) did not include the "X"s. currently used as wildcards. Instead, the last digits represented the volume. For example, on Magellan, a volume_set_ID "USA_NASA_JPL_MG_0001" was used ONLY for the volume with volume_ID of "MG_0001". Subsequent volumes in the same set had volume_set_IDs that differed in the final field.

If a set of volumes was to be distributed as one logical unit, the volume set ID included the range of volume IDs.

EXAMPLE

USA_NASA_PDS_VG_0001_TO_VG_0003 for the three volumes that comprise the Voyager Uranus volume set.

19.6 Logical Volume Naming

Logical volumes will retain the volume and volume set naming used at the physical volume level. For further information, see Appendix A, Volume Object.

19.7 Exceptions to This Standard

In some rare cases, machine or software restrictions may exist on volume IDs. Also, volumes made in the past may have IDs which do not meet this standard and there may be compelling reasons for keeping the same volume ID when making a new copy of the data. All new data sets, however, should use this standard.